

BL

PAT-NO: JP361285570A
DOCUMENT-IDENTIFIER: JP 61285570 A
TITLE: VOICE INPUT DEVICE

PUBN-DATE: December 16, 1986

INVENTOR-INFORMATION:

NAME COUNTRY
OSHIMA, YOSHIMITSU
AGO, MASAHIRO
YABUUCHI, SHIGERU
KITAHARA, YOSHINORI

ASSIGNEE-INFORMATION:

NAME COUNTRY
HITACHI LTD N/A

APPL-NO: JP60126029
APPL-DATE: June 12, 1985

INT-CL G06 F 015/20 , G06 F 003/16 ; G10 L 003/00 , G10 L
(IPC): 003/00

ABSTRACT:

PURPOSE: To attain the automatic input of punctuation marks by utilizing the intervals of speaking, a sentence intonation, parts of speech and conjugation of words, a sentence structure and meaning information respectively to estimate the positions of punctuation marks.

CONSTITUTION: An input voice 5 is analyzed at a phoneme recognition part 11 of a voice recognition part 1 and analyzed to the phoneme information. Thus a series 111 of KANA (Japanese syllabary) codes are delivered and the basic frequency 121 of the voice 5 is extracted by a pitch extracting part 12. While the power 131 of the voice 5 is detected by a power detection part 13. A voice data

processing part 14 receives the outputs 111□131 of each part to perform the data processing and delivers the KANA character strings including the punctuation mark candidates. A morpheme analysis part 2 divides an input KANA character string for each morpheme. A sentence structure meaning analysis part 3 uses the output of the part 2 as an input and performs analysis part 3 uses the output of the part 2 as an input and performs analysis by means of the structure and the meaning of an input sentence to obtain the relationship among component clauses. Then a punctuation mark processing part 4 estimates the positions of punctuation marks based on the result of analysis of the part 3 and applies them automatically.

COPYRIGHT: (C)1986,JPO&Japio

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A)

昭61-285570

⑬ Int.Cl.

識別記号

庁内整理番号

⑭ 公開 昭和61年(1986)12月16日

G 06 F 15/20

A-7010-5B

3/16

7341-5B

G 10 L 3/00

Z-8221-5D

GLA

審査請求 未請求 発明の数 1 (全9頁)

⑮ 発明の名称 音声入力装置

⑯ 特 願 昭60-126029

⑰ 出 願 昭60(1985)6月12日

⑱ 発 明 者 大 島 義 光 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内

⑲ 発 明 者 阿 呉 正 博 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内

⑳ 発 明 者 荻 内 繁 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内

㉑ 発 明 者 北 原 義 典 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内

㉒ 出 願 人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地

㉓ 代 理 人 弁理士 小川 勝男 外1名

明 細 書

発明の名称 音声入力装置

特許請求の範囲

1. 音声を用いた文章入力装置において、発声の間、文の抑揚、語の品詞または活用、および文の構文、意味情報を利用して句読点を打つべき位置を推定し、句読点を自動的に入力する音声入力装置。

発明の詳細な説明

(発明の利用分野)

本発明は音声を利用した文章入力装置、特に音声による句読点入力方式に関する。

(発明の背景)

文章をキーボードを用いて入力する場合は、通常キーボード上に句点、読点のキーが用意されているので、これまで問題となることはなかった。これに対して、通常の会話音声では句読点に対応するものがないので、音声によつて文章を入力する場合、新たに句読点の入力方式を設定する必要が生じてくる。

一つの方法として、読点に対して「テン」、句点に対して「マル」と発声して入力する方法があるが、意識的に行なわなくてはならず、特に原稿なしで音声によつて直接文章する場合には、入力作業者の思考を乱すので、適切ではない。

(発明の目的)

本発明は、このような句読点入力のための不自然さを解消し、音声による自然な文章入力方式を提供することを目的とし、具体的には、発声の間、声調、前後の構文・意味などの情報を利用して、自動的に句読点を入力する方式を提供することを目的とする。

(発明の概要)

句点と読点に分けて説明する。

「。」は文末につけるのが原則である。そして文とは一定の意味的なまとまりを持った陳述であり、音響的に見ると一定の特徴を持っている。

まず、音声のポーズ(無音区間)の観点で見ると、文内でも文節の境界を示すポーズが現われるが、ポーズ長は文節間の意味的な結合により決定

され、結合が弱いほどポーズ長は長くなる。そして最も弱い結合の場合はポーズ長(10)が経験的に300ms以上になるとされている(参考文献: 箱田他「文音声のポーズ挿入規則」, 日本音響学会音声研究会資料, 574-64, 1975年)。そこでまず、300ms以上のポーズがあれば、そこが文末の候補になる。

次に、イントネーション(抑揚)の観点で見る。イントネーションは音声の基本周波数によって表現される。基本周波数から単語アクセントなどによるアクセント成分を除いたものを考えると、これは、文頭で立上り、その後文末に向って緩やかに低くなつていき、ある一定の周波数に近づく。(参考文献: 日本放送協会編「日本語発音アクセント辞典」解説の部「共通語の発音とアクセント」第3章)。これを図示すると、第1図のようになる。収束の極限の周波数(F_{min})は、話者の最低周波数と考えられるので、事前に話者の最低周波数を音声認識部に記憶しておいて、この周波数と入力音声の基本周波数を比べることにより、文末

に近いところにいるかどうかを推定することができる。第1図にはポーズの例も示してあるが、ポーズとイントネーションの情報を併用すれば、処理の精度をより高くすることができる。

以上の音響情報だけでは、しかし、確実な処理は望めず、文の途中で似たような音響パターンが現れた場合には、判定を誤る可能性がある。これを補償するためには、文の持つ言語的情報も併用する必要がある。

日本語の文では、特殊な表現(例置法による表現など)を除くと、文末の形態を品詞ないし活用の観点で分類すると次のようになる。

- (1) 用言(動詞、形容詞、形容動詞)、助動詞、補助用言(補助動詞、補助形容詞)の終止形。
- (2) 同じく命令形。
- (3) 終助詞。
- (4) 名詞単独。(体言止めの場合)
- (5) サ変動詞語幹、形容動詞語幹。

このうち圧倒的に多く現れるのが(1)である。

(2)は話言葉的な表現であり、小説などで用い

られるか、「少年よ大志を抱け」などの格言で用いられるのみである。(3)は通常の文章でも疑問文や反語的な表現で用いられる。「か」など)。(4)も通常の文章で用いられるが、「春は曙」「花はさくらぎ」「酒は醒」など限られた表現で用いられるが、新聞などで限られた紙面にできるだけたくさんの記事をつめ込むために用いられるのみである。通常の文章では、タイトル、箇条書き文の中で用いられることもあるが、数は少ない。(5)も上と同様の理由で新聞などで用いられるが、通常の文章で用いられることは少ない。したがって、通常の文章では、(1)(3)についての処理を用意しておけばよい。

(3)については、入力文の形態素解析を行なつて、それが終助詞であることを認識できれば、それにより文末であることが判別できるので、「。」を打つことができる。

(1)の場合、現代の日本語では問題点が一つある。すなわち、現代語では、形容動詞と一部の助動詞(「だ」「ようだ」「そうだ」「みたいだ」)

を除いて、終止形と連体形が同形なので、活用形だけでは文末であるか否かを判定できない。これを解決するには、前記の音響情報による処理を併用すればいいと考えられる(連体形は通常後続の名詞を修飾するのに用いられるので、後続の名詞と一緒に発声されることが多い。)が、次のように入力文の持つ構文や意味の情報を利用すれば、より精度高く[.]を打つ位置を判定することができる。

まず、用言の直後に、通常文頭に置いて用いられる接続詞(「そこで」「したがって」「すると」など)が来ている場合、その用言は終止形であると判定することができる。したがって、その用言の直後に[.]を付与することができる。

次に用言が連体形の時、後続の名詞を修飾するわけであるが、この場合、その名詞は用言の格になっている場合が多い。そこで、入力文の構文意味解析を行なつて用言に係っている名詞文節(連用名詞文節)と用言が連体修飾している名詞文節の格をチェツクし、連体修飾されている名詞

文節の格に対応するものが連用名詞文節でない場合は、その名詞が用言に連体修飾されている推定することができ、したがって用言は終止形ではなく連体形であると推定することができる。このことは連用修飾側の名詞文節にその用言の必須格にあたるものがなく、連体修飾されている名詞が欠けている必須格の対応している場合には、推定の精度がより高くなる。逆に、構文的には連体修飾されていると仮定できる名詞に対応する格が用言の格フレーム中になく、その用言は名詞を連体修飾しているのではない。すなわち、用言は連体形ではなく終止形であると推定することができる。

なお、前記(2)の活用語の命令形の場合、終止形と違って紛れがないので、簡単に検知でき、文末であると判定することができる。

次に、読点(「,」:テン)について説明する。読点については、それがあつた場合とない場合とについて音学的に明確な差を認めることは難しい。そこで読点の付与は主に構文的に決定することに

なる。

日本語では読点のうちかたについて、英語のコンマのように明確なルールは設定されていないが、通常よく読点がつたれる部分を箇条書きにしてみると、次のようになる。

(1) 複文・重文の境界

例:「～し、～した。」

「～したので、～した。」

(2) 並立表現

例:「りんご、みかん、およびバナナ」

(3) 飛び越し係りの文節

直後の用言を越えてより後方の文節にかかる連用修飾文節など。

例:「渡辺刑事は、逃げ出した賊を追いかけた。」

この場合、「渡辺刑事は」は「逃げ出した」に係るのではなく、「追かけた」に係る。

(4) 文と文または節と節を結ぶ接続詞のうしろ

例:「しかし」「そこで」「さらに」

以上の場所に読点を付与すれば、通常の文章で

は大むね妥当な文章が得られると考えられる。

次に各々の場合、の自動付与の方法について述べる。

まず(1)の場合複文については用言のうしろの接続助詞を検知することにより簡単に判定できる。重文の場合は用言が連用形(連用中止の形)かうしろに接続助詞「て」がついた場合であるから、この場合も判定は簡単にできる。

(2)の場合、通常の文の他の名詞文節(連用文節、連体文節)には助詞がついているのが普通であるが、並立表現中の名詞文節には助詞がついていないので、これを検知することにより判定することができる。ただし、「インドりんご」のように複合語となる場合にも名詞に助詞がつかないのて区別が問題となるが、これは、発音の間を調べ、並立表現では名詞と名詞(または接続詞)のあいだで一拍間が置かれるのに対し、複合語では両者のあいだに間が置かれず続けて発音されることを利用して、判別することができる。

(3)の場合、入力文の構文意味解析を行ない、

各文節の係り受け関係を明らかにすることにより、「テン」をうつべきか否かを判定することができる。

(4)の場合、接続詞を検知することにより、簡単に判定できる。

【発明の実施例】

以下、本発明の実施例により説明する。

第2図に本発明の一実施例を示す。第2図で、1は音声認識部、2は形態素解析部、3は構文意味解析部、4は句読点処理部である。1の音声認識部より順に各部の動作を説明する。

第3図に音声認識部1の内部構成を示す。11は音声認識部で入力音声5を解析して、音韻情報に分解し、カナ文字コードの系列111を出力する。音韻認識部の詳細はここでは示さないが、例えば、伊藤部達著「音声タイプライタの設計」CQ出版社発行(1983年)の第2章、第3章記載の内容にのっとり構成することができる。

12はピンチ抽出部で、入力音声の基本周波数を抽出する。その出力121は、第1図上段のグ

ラフのようになる。このピッチ抽出部は、例えば、新英康永著「音声認識」共立出版発行（1978年）の第2章2、4節に記されている音声の分析法、同じく2、6節Bに記されているピッチ抽出法を用い、さらにアクセント保存の要素除去には広瀬啓吉・藤野博也・山口幹雄「基本周波数パタンの特徴の自動抽出」日本音響学会音声研究会資料S83-35（1983年）のアクセント推定アルゴリズム、フレーズ推定アルゴリズムを利用することにより、構成することができる。具体的にはアナログデジタル変換器、マイクロプロセッサ、読出し専用メモリ、ランダムアクセスメモリなどを用い、上記の各アルゴリズムを読出し専用メモリに格納したものとして構成される。

13はパワー検出部で、入力音声のパワーを検出する。その出力131は、第1図下段のグラフのようになる。パワー検出部13の詳細はここでは示さないが、既知の整流回路および積分回路などで構成することができる。

14は音声データ処理部で、各部の出力111

～131を受けてデータ処理を行ない、区切り記号（の候補）を含んだカナ文字列を出力する。この音声データ処理部は、市販のマイクロコンピュータ、ランダムアクセスメモリ、およびアナログデジタル変換回路などを用いて構成することができる。

第4図に、音声データ処理部14における句点情報抽出のための処理フローを示す。第A1ステップ、第A2ステップで音声入力の基本周波数データとパワーを監視し、F₀にはパワーが0となる直前の基本周波数が保持されるようにする。音声入力のパワーが0となつたら次へ進む。パワー0の時間をはかる（第A3ステップ）。そして、パワー0の時間長が300μsec以上であれば（第A4ステップ）、第A5ステップへ進む。第A5ステップでは、パワーが0となる直前の周波数F₀と話者の最低周波数F_{min}との差（ΔF）をとり、これがある一定値S以下ならば（第A6ステップ）、文末であると見なし、句点候補を出力する。この句点候補のデータは音韻情報（カナコ

ード列）と混合されて、出力線6に出力される。出力線6上のデータの形式を第5図に示す。「。」で示してあるのが句点候補である。ここでカギカッコをつけているのは、句点の候補の意を示すためである（以下同様）。

なお、話者の声の最低周波数F_{min}は、予じめ登録しておくか、前記音声データ処理装置内に学習機能を設け、第4図のフローチャートで句点候補を出力したときの基本周波数F₀をF_{min}として保存し、F_{min}より低いF₀が入力されたときはこれを更新するというような方法で設定することが可能である。

また、閾値Sは、実験的に適切な値に設定しておく。

なお、第3図において、ピッチ抽出部12、パワー検出部13、音声データ処理部14を独立の回路としているが、音声認識内でもそれぞれ12～14と類似の回路が必要であり、両者を兼用することも可能である。

次に第2図の形態素解析部2は、カナ文字列と

句点情報からなる音声認識部の出力6を受けとり、形態素解析を行なう。形態素解析は単語辞書および形態素列の連接可否を記した接続表などを用いて、入力カナ文字列を形態素単位に分割する処理である。形態素解析部の具体的な構成法に関しては、相沢・江原「計算機によるカナ漢字変換」NHK技術研究第25巻第5号頁23～60（1973年）、あるいは特開昭59-44025などを参照することにより構成可能である。第5図のカナ文字列入力に対応する形態素解析部2の処理結果を第6図に示す。各形態素の読み、表記、品詞、活用、および形態素解析は文節の区切りも認識するので、文節の位置にマークを記している。

構文意味解析部3は、上記の形態素解析部2の出力7を入力として、入力文の構文と意味を利用した解析を行ない、構成文節の係り受け関係をもとめる。構文意味解析は格フレーム辞書、日本語の構文規則を利用して係り受け関係をもとめる処理で、具体的には発明特開59-11918記載の内容などを参照して構成することができる。

「ワタシハ サクラガ サイタノデ コウエンヘ
デカケタ」という音声入力が音声認識部、形態素
解析部、構文意味解析部を次々に通過してくると、
第7図に示すような処理結果が得られる。」第6
図の品詞、活用および、文節始の情報に加えて、
係り先の情報が付加されている。係り先欄の数字
は、表の左端の形態素に付与した数字に対応して
いる。「なし」とあるのは、構文意味解析で文末
と認定した文節である。

句読点処理部は、第7図のような処理結果をも
とに句読点の位置を推定し、自動的に付与する。
句読点抽出部は、具体的には、市販のマイクロプ
ロセッサ、メモリ等により構成することができる。

次に、句読点処理部4の処理内容を句点、読点
それぞれに分けて説明する。

第8図に、音声認識部1で出力された句点候補
のデータを、形態素解析部2、構文意味解析部3
の処理結果を利用して判定するための処理手順を
示す。

第B1ステップおよび第B2ステップで、構文

意味解析部3の出力8を走査し、句点候補のデー
タを探す。句点候補が見つかったら(第7図では
14行目の「。」印)、次に、さらにその次の文
節のデータが、文頭に位置して用いられる接続詞
ならば、その手前が文の境界と推定されるので、
句点候補を句点と認定して出力する(第B3ステ
ップ、第B4ステップ)。

次の文節が接続詞でないならば、次に、直前の
文節の文節末の形態をチェックする(第B5ステ
ップ、第B6ステップ)。その品詞が活用語か否
かをチェックし、活用語でなければ、次に終助詞
かどうかチェックし、終助詞ならば文の終りと推
定し、句点候補を句点と認定し出力する(第B7
ステップ、第B8ステップ)。さもないければ何も
せず(句点候補を句点の認定せず、無視する)、
次の句点候補探索ループへ再度進む。

直前の文節末が活用語ならば、さらにそれが形
容動詞または形容動詞型活用の助動詞かどうかを
チェックする(第B9ステップ)。そうであれば、
その活用形をチェックし、活用形が終止形ならば

文の終りと判定できるので、句点候補を句点と認
定し出力する(第B10ステップ、第B11ステ
ップ)。終止形でなければ、文の終りでないので、
何もしない。

直前の文節末が、形容動詞または形容動詞型活
用の助動詞以外の活用語ならば活用形をチェック
し(第B12ステップ)、終止形(または連体形)
ならば、それが連体形か終止形かを判定するた
めに構文意味解析部3の処理結果8の係り先欄(第
7図参照)を参照する(第B15、B16ステッ
プ)。係り先欄に係り先なしと記されているれば、
そこが文の終りであると推定できるので、句点候
補を句点と認定し出力する(第B17ステップ)。
さもないければ何もしない。

以上の処理により、通常の多くの文で句点を扱
うべき位置を検知し、付与することができる。

次に読点の処理について説明する。

第9図に、句読点処理部4における読点の処理
手順を示す。

第C1ステップで構文意味解析部3の出力8を

走査し、文節自立語の品詞を次々にチェックする
(第C2ステップ)。文節の自立語が用言ならば
第C3ステップに進み、文節末の形態をチェック
する。それが活用語の連用形部尾(助動詞、補助
用言の連用形を含む)、あるいは接続助詞ならば、
重文または複文の境界と推定されるので、読点を
付与する(第C4ステップ)。連体形または終止
形部尾の場合は前記の句点処理時にチェックされ
ているが、終止形と認定されれば句点を付与され
ており、読点を付与する必要はない。連体形の場
合は通常後続の名詞と一続きに表現されるので、
読点を付与する必要はない。

次に、文節の自立語が接続詞ならば、それが文
と文、または節と節をつなぐ接続詞(「しかし」
「そして」「さらに」など)であるかどうかをチ
ェックする(第C5ステップ)。そうであれば文
節のうしろに読点を付与する(第C6ステップ)。
そうでなければ(この場合は、「または」「およ
び」「あるいは」などの語と語、または句と句を
結ぶ接続詞)、何もしない。

文節の自立語が名詞ならば、文節の付属語をチェックする(第C7ステップ)。付属語に断定の助動詞(「だ」「である」「です」)またはそれに準じて名詞に直接接続する助動詞(「らしい」「みたいだ」)がある場合、文節全体としては用言文節と同等な性格を持つので、第C3ステップにもどり、用言文節と同じ処理を行なう。

文節の自立語が名詞単独で、付属語が付いていなければ、それは他の名詞と並立関係にある名詞か複合語の要素であるかどうかである。これを区別するためには後続の節とのあいだで発声の間があるかどうか調べればよいが、これには最初に説明と音声認識部1での処理中、句点候補として取り上げられたものの以外の発声の間を、読点候補としてデータ中に入れておけばよい(第10図参照)。[、]が読点候補。これを利用して読点候補の有無を調べ(第C8ステップ)、あれば並立文節と認定して読点を出力し、なければ複合語要素と見なして何もしない。

名詞に助動詞がついているならば、次に、その文

節の係りの型をチェックする(第C10ステップ)。名詞文節の係り型は文節末の助動詞をチェックすることに判定することができる。格助詞「の」ならば連体修飾、その他の格助詞および係り助詞、副助詞ならば、連用修飾である。文節の係り型が連体修飾のときは通常後続の名詞と一緒に表現されるので何もしない。連用修飾の場合はさらにその係り先のチェックを行なう(第C11ステップ)。係り先が近く(直後)の用言ならば何せず、直後の用言を越えてより後方の用言に係る場合は、文節のうしろに読点を付与する(第C12ステップ)。これを例で説明すると、例えば第7図の2番目の文節「桜が」は直後の文節「咲いたので」に係っているので読点を付さず、第1番目の文節「私は」は「咲いたので」を越えて「出かけた」に係っているので、読点を付与する。なおここで、直後の用言とは、用言として直後という意味であり、注目している名詞文節と用言の間に他の名詞文節や副詞文節が入ることを妨げない。ちなみに第7図の例では、接続助詞「ので」のうし

ろにも読点が付与される。

第7図に対し句読点処理を施した最終結果を第11図に示す。

なお、以上の発明の概要および実施例の説明において、音声認識部1、形態素解析部2、構文意味解析部3などの処理データをすべて利用して句読点の付与位置を判定する方法について説明したが、これを部分的に利用して、例えば音声認識1のみのデータ、あるいは音声認識部1と形態素解析部2の処理結果を利用して句読点の位置を判定するような構成にすることも可能である。

(発明の効果)

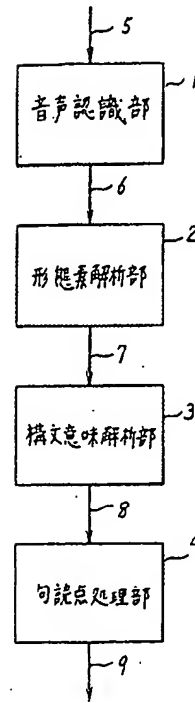
以上本発明によれば、音声入力された文章に自動的に句読点を付与することが可能となる。したがって、本発明に基づいた音声入力装置を用いて文章を入力作業者は、句読点の入力について気を遣らす必要はなくなり、自然な調子で入力することが可能となるので、音声による文章入力の効率向上に寄与することもできる。

図面の簡単な説明

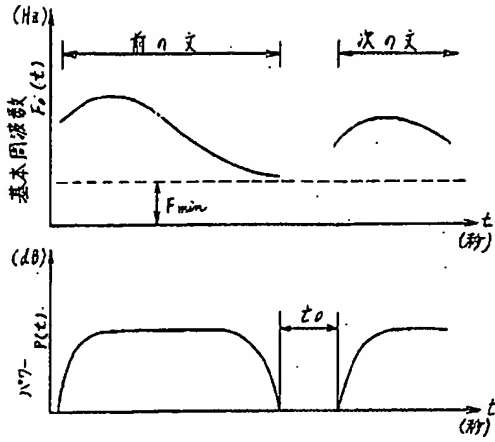
第1図は音声入力による文と文の境界における音声の基本周波数とパワーの変化を示した図、第2図は、本発明における実施例の全体構成を示す図、第3図は第2図の音声認識部1の詳細を示した図、第4図は第3図の音声データ処理部14における句点候補判定手順を示すフローチャート、第5図は音声認識部1の処理結果の例、第6図は同じく形態素解析部2の処理結果の例、第7図は構文意味解析部3の処理結果の例である。第8図は句読点処理部4における句点処理手順を示すフローチャート、第9図は読点処理手順を示すフローチャートである。第10図は、並立文節があるときの句読点処理部4への入力、第11図は第7図のデータに正しく句読点を施した結果である。

代理人 井理士 小川勝男

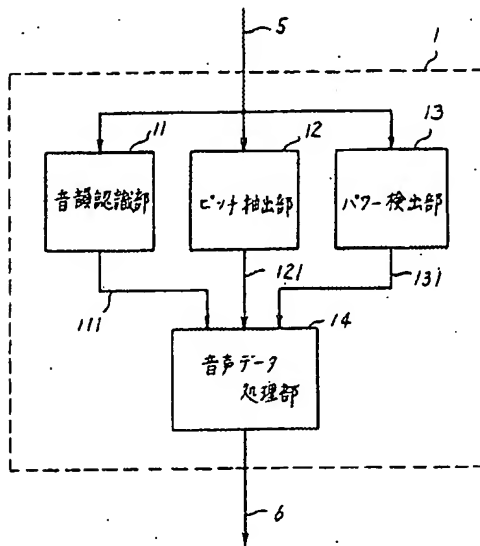
第 2 図



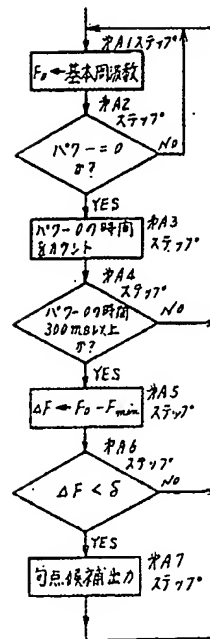
第 1 図



第 3 図



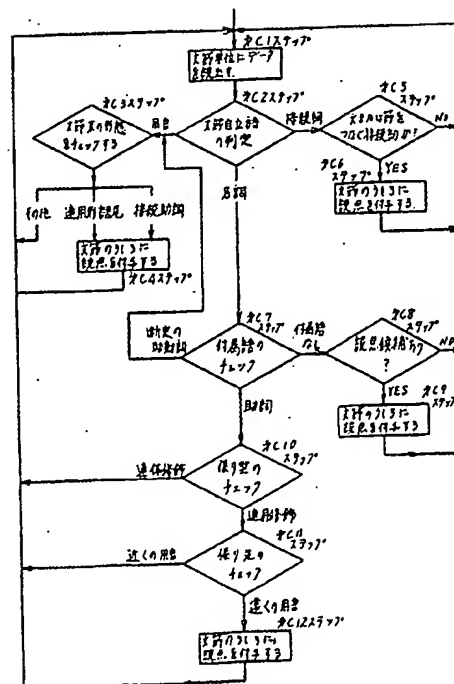
第 4 図



第 6 回

| 番号 | 形態素 | 文節端 | 品詞 | 活用 | 係り先 |
|----|-----|-----|------|------|-----|
| 1 | 禾ム | | 名詞 | | 11 |
| 2 | は | ★ | 係助詞 | | |
| 3 | 桜 | | 名詞 | | 5 |
| 4 | が | ★ | 格助詞 | | |
| 5 | 咲 | | 動詞 | 加行Ⅱ段 | 11 |
| 6 | い | | 動詞語尾 | 連用形 | |
| 7 | た | | 助動詞 | 連体形 | |
| 8 | ので | ★ | 接続助詞 | | |
| 9 | 公園 | | 名詞 | | 11 |
| 10 | へ | ★ | 格助詞 | | |
| 11 | 出か | | 動詞 | 下段 | なし |
| 12 | け | | 動詞語尾 | 連用形 | |
| 13 | た | ★ | 助動詞 | 終止形 | |
| 14 | 〔。〕 | | | | |

第 9 回



第 10 図

| 番号 | 形態素 | 節端 | 品詞 | 活用 | 係り先 |
|----|-----|----|-----|----|-----|
| 1 | りんご | ★ | 名詞 | | |
| 2 | 【、】 | | | | |
| 3 | みかん | ★ | 名詞 | | |
| 4 | 【、】 | | | | |
| | およ | ★ | 接続詞 | | |
| | バナナ | ★ | 名詞 | | |
| | | | | | |

第 11 図

私は、桜が咲いたので、公園へ出かけた。